

人工智慧

1

— 語言翻譯

指導老師：吳智鴻 老師

學生：ACS096108 李佩珊
ACS096141 謝昀臻

ACS096122 陳藝方
ACS096144 羅敏華

大綱

2

- 原理簡介-----陳藝方、謝昀臻
- 相關軟體-----李佩珊
- 目前應用-----羅敏華

歷史發展

3

- 翻譯相關研究在17世紀就開始發展，哲學家笛卡兒提出所謂的「世界語言」（universal language）概念，試圖編寫類似於統一語言的辭典
- 在1954年喬治城大學實驗將超過六十個俄國句子全自動機械的翻譯成英語。被視為機器翻譯實際可行的開端。然而，ALAPC（自動語言處理顧問委員會，Automatic Language Processing Advisory Committee）在1966年提出的一項報告中表明十年來的機器翻譯研究進度緩慢，未達預期
- 直到近1980年代，由於電腦運算科技的進步，以及相對的演算成本的降低，才使得政府與民間企業針對機器翻譯再次有了興趣，特別是在統計法機器翻譯的領域上

機器翻譯

4

- Machine Translation (MT)
- 是利用計算機把一種自然源語言(source language)轉變為另一種自然目標語言(target language)的過程，一般指自然語言之間句子和全文的翻譯

原理技術

5

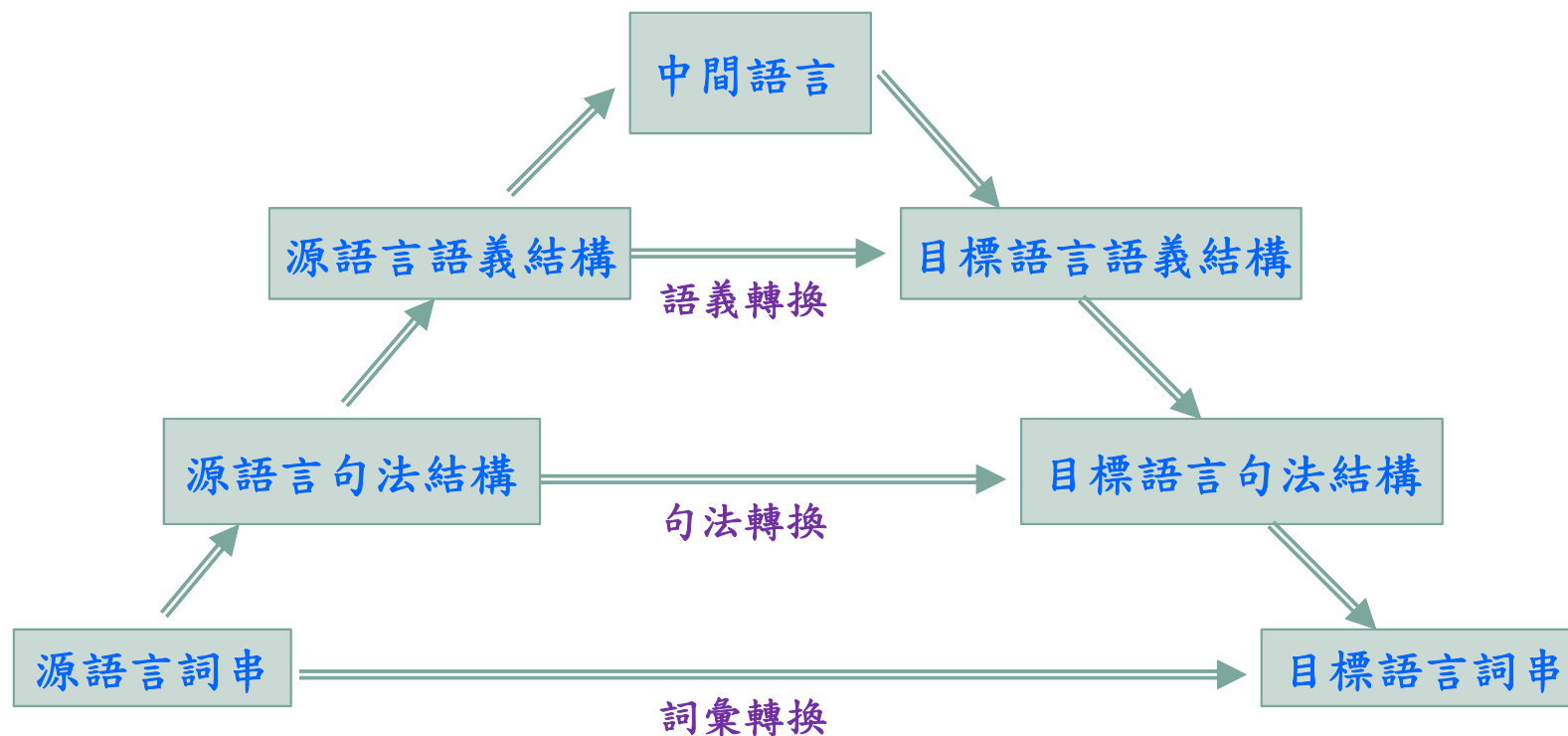
- 基於規則(Rule-based)
 - 轉換法(Transfer)
 - 中間語言(Interlingua)

- 基於語料庫(Corpus-based)
 - 基於實例(Example-based)
 - 基於統計(Statistic-based)

基於規則

6

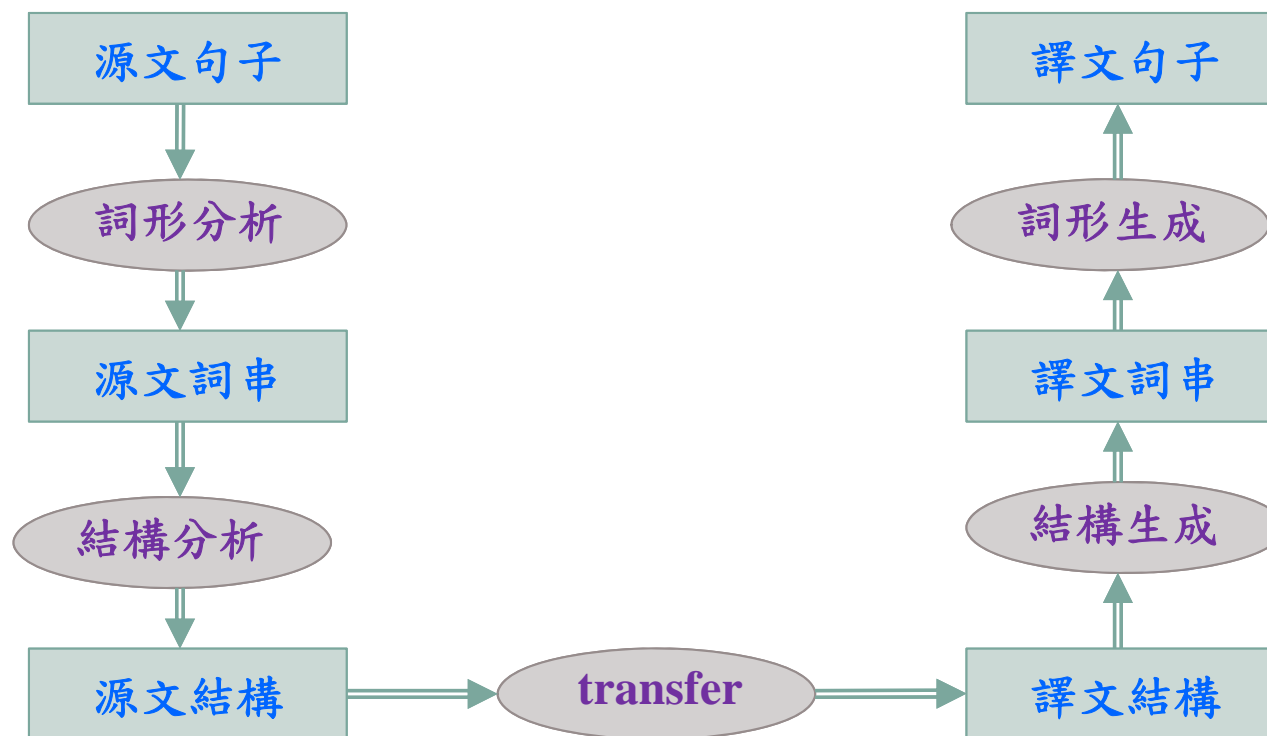
□ 辭典+規則庫



基於規則(Cont.)

7

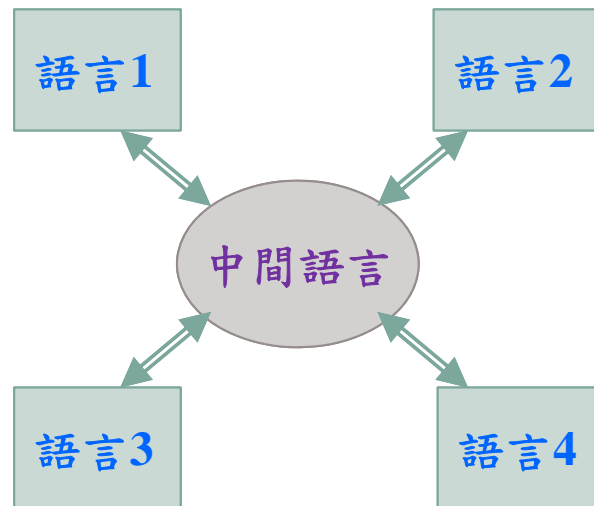
□ 轉換法(transfer)



基於規則(Cont.)

8

- 中間語言(Interlingua)
 - ▣ 自然語言的電腦表示形式的系統化，試圖創造出一種獨立於各種自然語言，同時又能表示各種自然語的人工語言。



Ex: 世界語、人工定義之語言

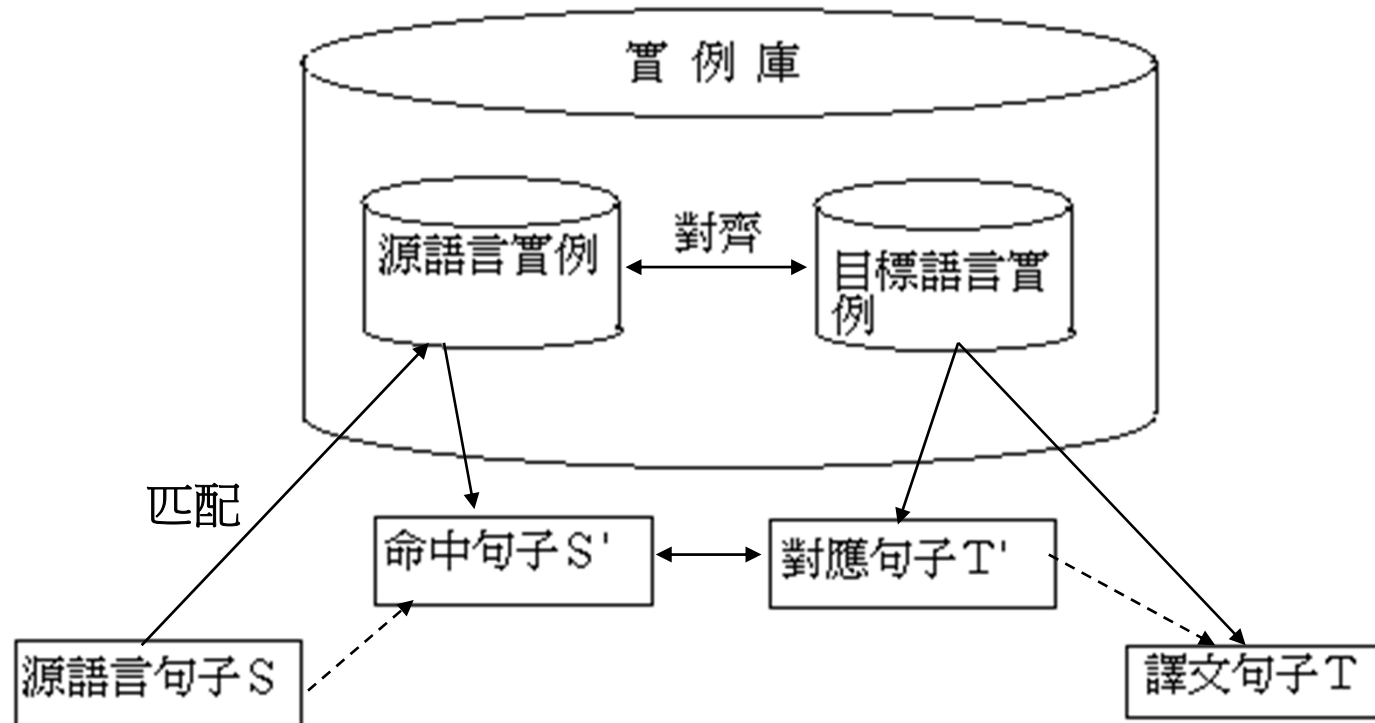
基於語料庫

- 機器翻譯系統的語料庫通常稱為平行語料庫 (Parallel Corpus)。平行語料庫當中記載了來源語言與目標語言間的對應關係。像是文章間的對應、句子間的對應、以及片語和詞彙上的對應關係。這些對應關係對統計式翻譯系統提供了一些統計上的參考與指導原則。

基於語料庫(Cont.)

10

□ 基於實例(Example-based)



基於語料庫(Cont.)

11

□ 基於實例(Example-based)

□ Ex:

英文實例	中文翻譯
He eats meat.	他吃肉
Acid eats metal.	酸腐蝕鐵

I eat potatoes. 我吃馬鈴薯



基於語料庫(Cont.)

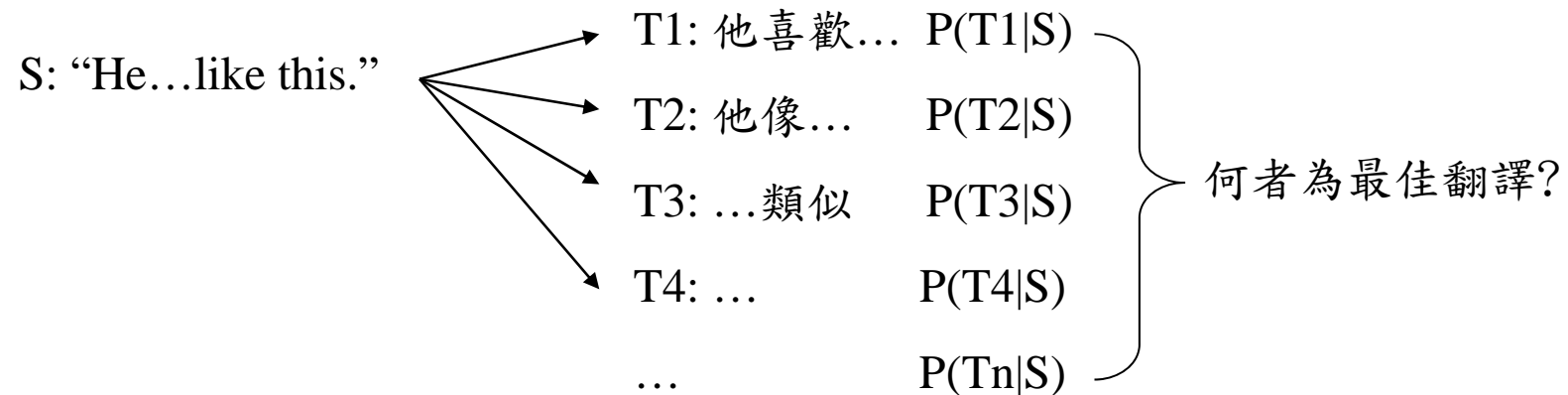
12

- 基於統計(Statistic-based)
 - 統計式的方法近幾年來受到相當大的重視，其原因主要是由於網路上大量資料的出現，使得語料的取得即為方便。這讓統計式翻譯得以先對大量資料進行統計後，給與每一條規則、詞彙與短語一個機率值。

基於語料庫(Cont.)

13

□ 範例：



※ $P(T|S)$: 譯文T對應到原文S的機率


Google翻譯

14

- 「Google 翻譯」是從數百萬份文件中搜尋句型，然後再判斷最佳的翻譯結果。
- 「Google 翻譯」採用真人翻譯的文件歸納句型規則，因此能挑選正確又傳神的翻譯。這個從大量文件內容中尋找句型的程序叫做「統計式機器翻譯」。
- 翻譯是由機器產生的，所以不太可能都完美無暇。「Google 翻譯」分析的真人翻譯文件越多，就越能產生優質的譯文。正因如此，翻譯的準確性會因語言配對而有所差異。



Google 翻譯

原文是 中文  翻譯成 英文 翻譯

一群豬哥

 聆聽

試試具有自動翻譯功能的最新瀏覽器。

下載 Google 瀏覽器

將中文翻譯為英文

A group of gentlemen

 聆聽

將 Google 翻譯用在我的：[搜尋](#) [影片](#) [電子郵件](#) [手機](#) [即時通訊](#) [工作](#)

[關於 Google 翻譯](#) [關閉即時翻譯](#) [隱私權](#) [說明](#)

相關軟體

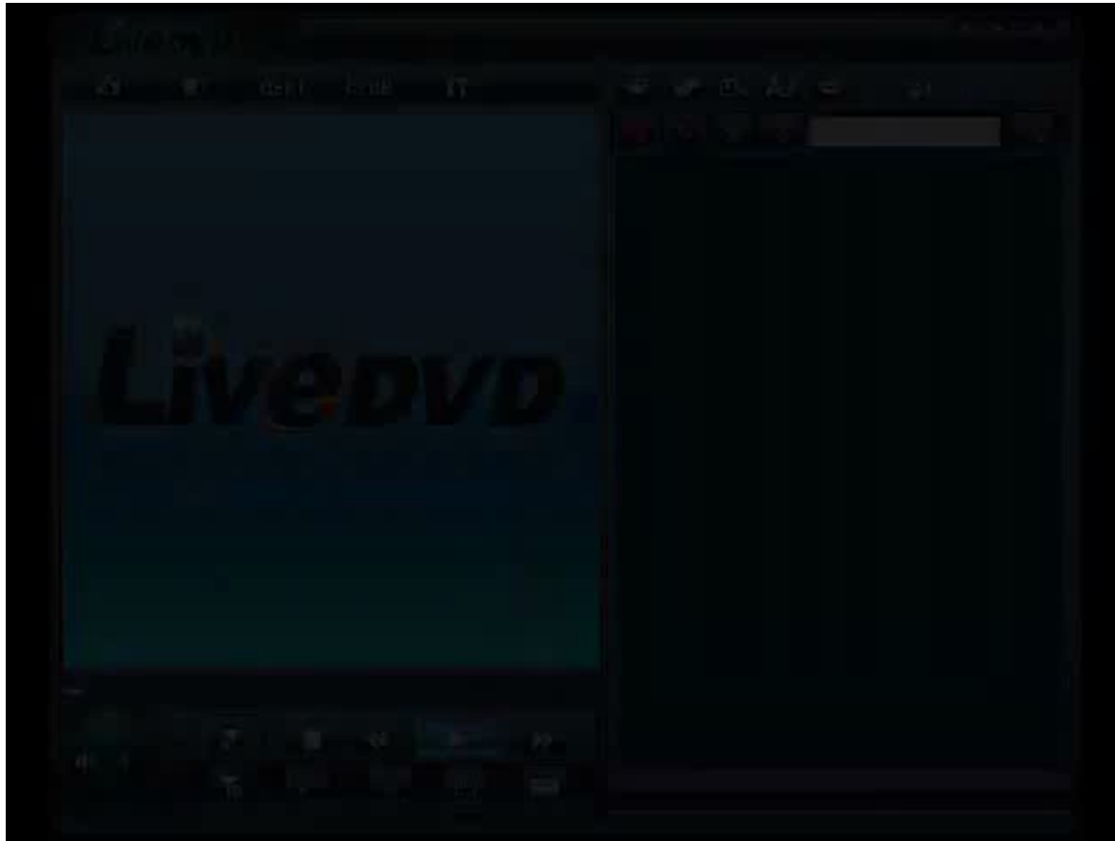
16

- 線上翻譯
 - [Google翻譯](#)、[Babelfish](#)、[bing翻譯](#)、[SYSTRAN](#)、[GramTrans](#) ...
- 網頁翻譯
 - Google Toolbar、Google Chrome
- 共享軟件
 - [巴比倫](#)、Dr.eye

目前應用

17

- LiveDVD
 - 能讓DVD電影成為學習各種語言的活教材



目前應用(Cont.)

18

□ LingoPal

- 包含多種情境，可用各種語言翻譯，不用怕到了國外卻無法溝通的窘境

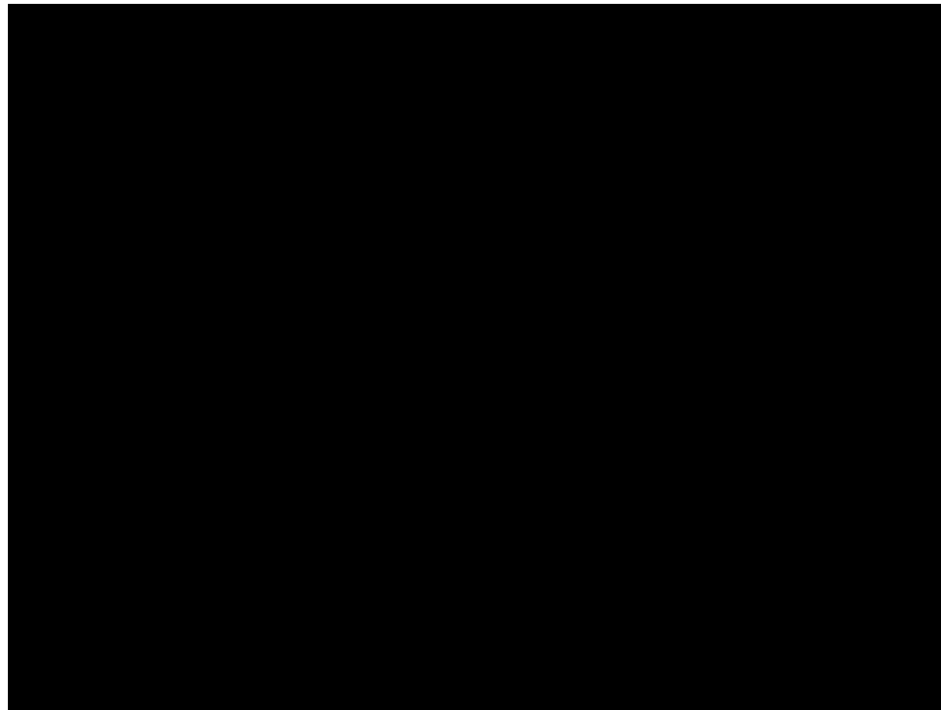


目前應用(Cont.)

19

□ Talkman

- 一款 PSP 即時語音翻譯與教學娛樂軟體，透過 PSP 專用 USB 外接麥克風，即可依照軟體中所預設的各種場合，進行各種語言的語音輸入與翻譯

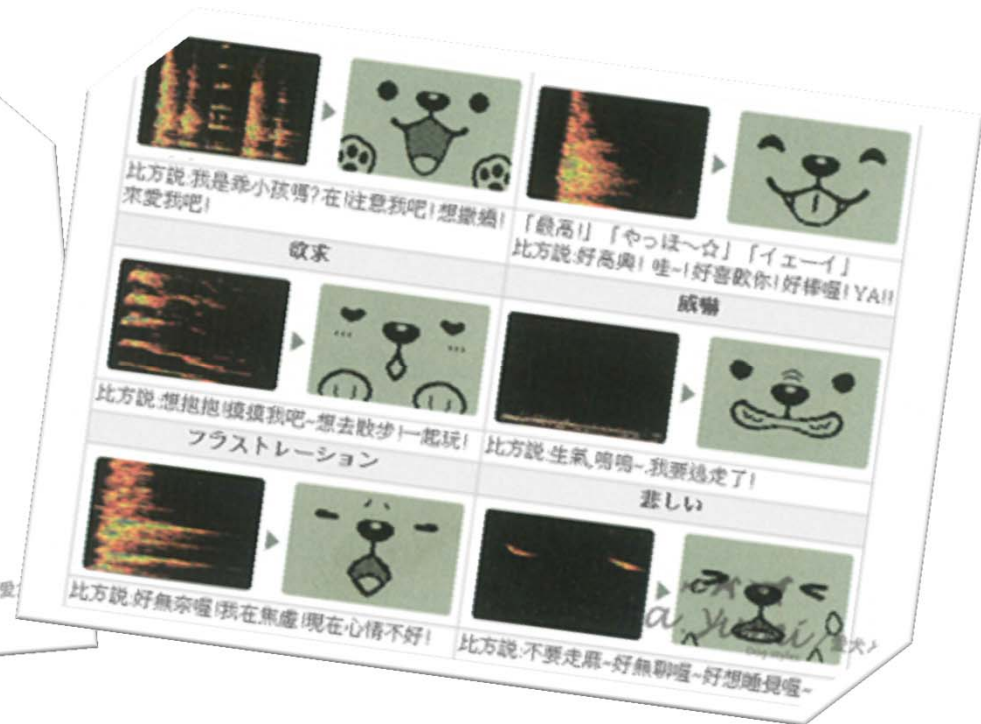


目前應用(Cont.)

20

□ Bowlingual

- 由日本 Takara 公司所發明，裝在狗的項圈上，透過無線麥克風把聲音傳到掌上型電腦，翻譯成六種狗的感情類別，以及兩百多句預設的"狗語"



目前應用(Cont.)

21



TAKARA
TOMY

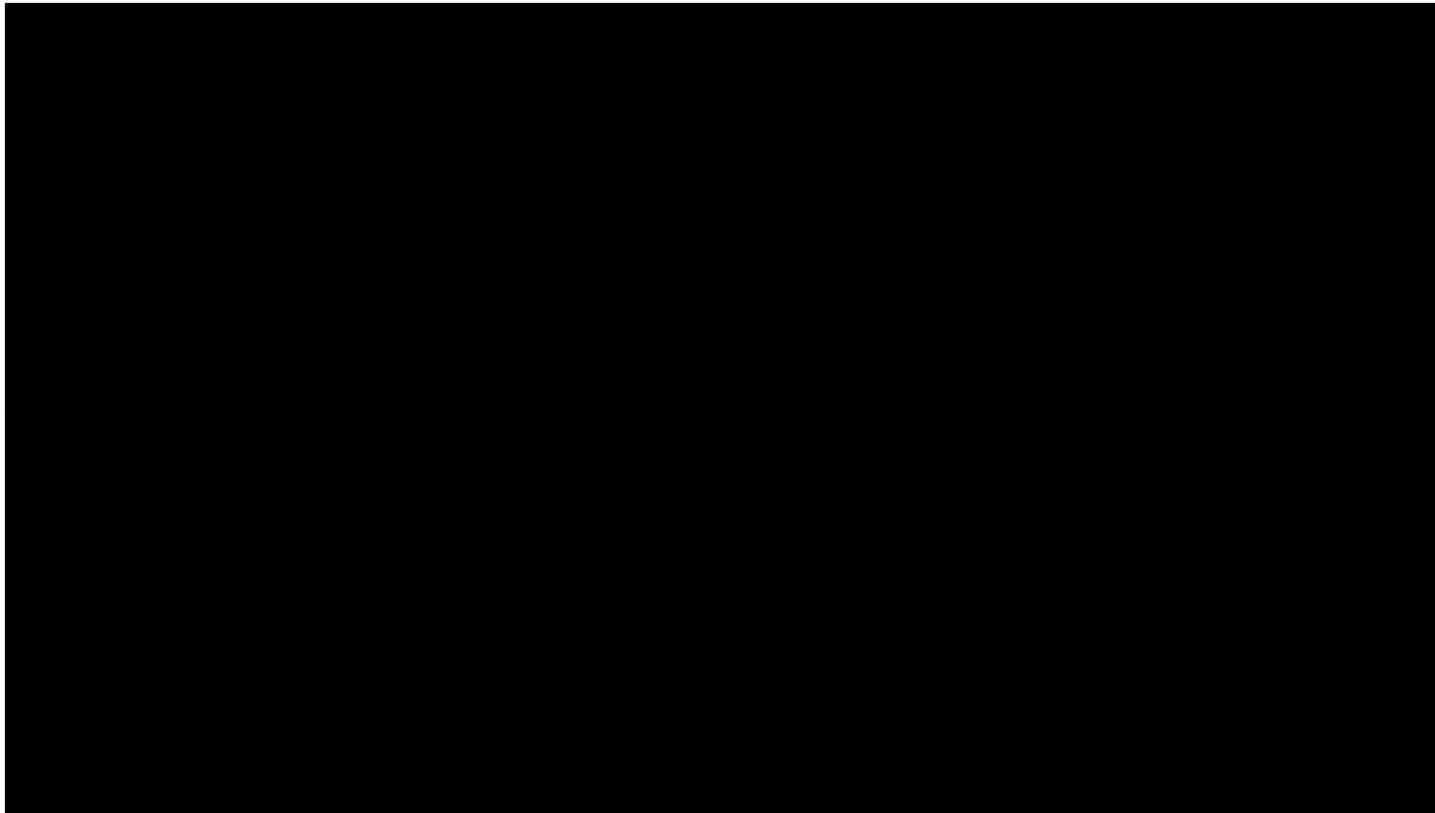
Bowlingual
バウリンガルボイス voice

目前應用(Cont.)

22

- Manslater

- 男人翻譯機---可以輕易聽懂女人話中的涵義



結論

- 語言是人們用來溝通與紀錄的重要工具，但不同的語言卻成為相互交流的障礙。一個良好的翻譯工具，使我們不僅能暢行無阻，又能加速學習與吸收各種知識。
- 機器語言發展至今仍不能超越專業的人腦，但對於我們這些初學者的幫助仍不小

參考資料

- World Lingo :
http://www.worldlingo.com/ma/enwiki/zh_tw/Machine_translation
- 陳鐘誠的網站:<http://ccckmit.wikidot.com/mt:smtbasics>
- 維基百科: <http://zh.wikipedia.org/zh-tw/%E6%9C%BA%E5%99%A8%E7%BF%BB%E8%AF%91>
- 百度百科:<http://baike.baidu.com/view/21352.htm>
- <http://203.64.42.21/course/2005/tgbcl/poko/ohcl-27.htm>